Practice of Epidemiology

Guideline-Based Physical Activity and Survival Among US Men With Nonmetastatic Prostate Cancer

Barbra A. Dickerman*, Edward Giovannucci, Claire H. Pernar, Lorelei A. Mucci, and Miguel A. Hernán

* Correspondence to Dr. Barbra A. Dickerman, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115 (e-mail: bad788@mail.harvard.edu).

Initially submitted August 14, 2018; accepted for publication November 27, 2018.

The survival impact of adhering to current physical activity guidelines after prostate cancer diagnosis is unknown. We therefore emulated a target trial of guideline-based physical activity interventions and 10-year survival among US men with nonmetastatic prostate cancer. We used observational data on 2,299 men in the Health Professionals Follow-up Study who were diagnosed with nonmetastatic prostate cancer from 1998 to 2010 and were free of conditions that might have precluded participation at baseline (first postdiagnostic questionnaire). We estimated their survival under several guideline-based physical activity interventions starting at baseline and ending at the development of conditions limiting physical ability. We adjusted for baseline and time-varying risk factors for death using the parametric g-formula. Compared with the observed 15.4% mortality risk, the estimated 10-year risks of mortality were 13.0% (95% confidence interval (CI): 10.9, 15.4) and 11.1% (95% CI: 8.7, 14.1) for ≥1.25 hours/week and ≥2.5 hours/week of vigorous activity, respectively, and 13.9% (95% CI: 12.0, 16.0) and 12.6% (95% CI: 10.6, 14.7) for ≥2.5 hours/week and ≥5 hours/week of moderate activity, respectively. We estimated that these men would have experienced clinically meaningful reductions in mortality had they followed current physical activity recommendations until the development of conditions limiting physical ability. These findings may help guide clinical recommendations for prostate cancer patients and the design of future randomized trials.

causal inference; cohort studies; lifestyle; parametric g-formula; physical activity; prostate cancer; survival

Abbreviations: BMI, body mass index; CI, confidence interval; HPFS, Health Professionals Follow-up Study; MET, metabolic equivalent of task.

Observational studies suggest that physical activity after prostate cancer diagnosis may decrease the risk of progression and mortality (1–5). Short-term randomized trials among prostate cancer survivors have demonstrated that physical activity interventions are feasible and improve body composition, biomarkers of cardiovascular health, and quality of life (6–8). To date, randomized trials have not reported on survival outcomes. To generate evidence supporting the incorporation of physical activity programs into routine cancer care, the American Society of Clinical Oncology recently called for large-scale trials of cancer survivors to evaluate the effect of physical activity interventions on outcomes such as survival (9).

The feasibility of a randomized trial of physical activity and survival among men with nonmetastatic prostate cancer is limited by the long follow-up that would be required for this slowly progressing disease. In the absence of data from such a trial, this effect needs to be estimated from observational studies (10). Previous observational studies of postdiagnosis physical activity and survival that did not adjust for prediagnosis activity have reported associations that may partially reflect the effect of lifelong physical activity habits rather than postdiagnosis changes in physical activity. Further, these studies did not report adjusted measures of absolute mortality and population attributable risks under realistic physical activity strategies. Therefore, the clinical impact of adhering to current guidelines for physical activity after prostate cancer diagnosis is unknown.

Here we estimated the effects on mortality of several physical activity interventions initiated at the time of prostate cancer diagnosis. To do so, we emulated a target trial of guideline-based physical activity interventions in men with nonmetastatic...
prostate cancer (11). We applied the parametric g-formula, a
generalization of standardization to time-varying exposures
and confounding, to estimate the 10-year risk of all-cause mor-
tality under several possible interventions.

METHODS

Target trial specification

The protocol of the target trial of physical activity inter-
ventions among men with nonmetastatic prostate cancer has
the following components (see also Table 1, available at

Eligibility criteria. Eligibility criteria include diagnosis with
nonmetastatic prostate cancer (tumor-node-metastasis (12) codes
T1–4/N0–1/M0) between 1998 and 2010, being 80 years of age
or younger, and having no history of: metastasis since diagnosis,
a recent cardiovascular event (myocardial infarction or stroke),
congestive heart failure, amyotrophic lateral sclerosis, or func-
tional impairment (difficulty climbing a flight of stairs or walking
8 blocks due to a physical impairment), under the assumption
that a physical activity intervention would not be feasible among
these men.

Outcome and follow-up. The outcome of interest is all-cause
mortality within 10 years of diagnosis. Each eligible man is fol-
lowed from baseline until death, incomplete follow-up, 10 years
after baseline, or the administrative end of follow-up in June
2014, whichever happens first.

Physical activity strategies. Eligible individuals are ran-
domly assigned to dynamic strategies of vigorous activity—at
least 1) 1.25 hours/week, 2) 2.50 hours/week, and 3) 3.75 hours/
week—and moderate activity—at least 4) 2.5 hours/week,
5) 5.0 hours/week, and 6) 7.5 hours/week—initiated at base-
line and continued over follow-up until the development of
conditions that may limit physical ability (functional impairment,
metastasis, myocardial infarction, stroke, congestive heart
failure, or amyotrophic lateral sclerosis). These are examples of
threshold interventions that impose a minimum threshold on activity (13). For example, at each follow-up time under
strategy 1, everyone engaging in <1.25 hours/week of vigorous
activity needs to increase their exercise level to ≥1.25 hours/
week of vigorous activity, and anyone engaging in ≥1.25 hours/
week of vigorous activity does not need to modify their exercise
levels.

The durations and intensities of these interventions are selected
to match current guidelines (6, 14) for physical activity among
cancer survivors: engaging in strategy 1 or 4 at a minimum
and engaging in strategy 2 or 5 for additional health benefits.
Higher levels of physical activity are explored through strat-
egies 3 and 6.

Activity is defined in units of hours/week rather than meta-
bolically equivalent of task (MET)-hours/week for consistency
with the guidelines, practicality for trial implementation, and
interpretabity for patients. Additionally, we assume that all
types of activity resulting in the recommended intensity and
duration will produce similar effects (15, 16).

Target trial emulation

We emulated the above target trial using data from the pro-
spective Health Professionals Follow-up Study (HPFS), an
ongoing cancer epidemiology cohort study of 51,529 US male
health professionals aged 40–75 years at enrollment in 1986.
Participants have reported detailed clinical and lifestyle inform-
ation at enrollment and every 2 years thereafter, including infor-
mation on weight, cigarette smoking, functional impairment,
and chronic diseases. A previous validation study in the HPFS
showed that self-reported and technician-measured weights
were highly correlated (Pearson’s r = 0.97) (17). Body mass
index (BMI) was calculated as weight (kg)/height (m)^2. Height
and parental history of myocardial infarction before age 60
years were ascertained in 1986.

Eligibility criteria. Men were asked on the biennial ques-
tionnaires whether prostate cancer had been diagnosed in the
prior 2 years. We verified the diagnosis (International Clas-
sification of Diseases, Ninth Revision, code 185) using medical
records and pathology reports from treating physicians and
hospitals. In a standardized review, we abstracted information
on diagnosis date, clinical stage, Gleason grade (18), prostate-
specific antigen levels, and initial treatment. We obtained
information on subsequent metastasis development from
prostate-cancer–specific biennial questionnaires sent to all
prostate cancer survivors and their physicians.

We applied all eligibility criteria to men in the HPFS and
also excluded men who did not have information on physical
activity, BMI, smoking history, primary treatment, clinical stage,
Gleason grade, or prostate-specific antigen level at diagnosis.
After these exclusions, 2,299 men were available for the anal-
ysis. Figure 1 shows a flow chart of participant selection.

Outcome and follow-up. Deaths were ascertained through
repeated mailings, telephone calls to nonrespondents, and searches
of the National Death Index. We defined baseline as the date
of return of the first postdiagnostic questionnaire and incom-
plete follow-up as questionnaire nonresponse.

Physical activity strategies. Physical activity has been re-
ported every 2 years since baseline using a validated question-
naire on type, frequency, and intensity of activity (19). Each
activity was assigned a MET value, which is a multiple of resting
energy expenditure. Vigorous activity was defined as activity
requiring ≥6.0 METs (running, jogging, bicycling, lap swim-
mee, tennis, squash/racquetball, calisthenics/rowing, other
aerobics). Moderate activity was defined as activity requiring
3.0–5.9 METs (walking, digging, moderate-to-heavy outdoor
activity, weight-lifting, low-intensity exercise). Men reported
the average number of hours per week spent in each activity
during the prior year. For a reported range of values, the mid-
point was used as a measure of weekly time spent in that activ-
ity. Time spent in activities was summed to obtain total weekly
duration of physical activity. In a validation study in the HPFS,
the correlation between questionnaire-based and diary-based activity
scores was 0.58 for vigorous activity (19). We truncated phy-
ical activity values at the 99th percentile to prevent implausible
values from affecting our analyses.

Statistical analysis

We used the parametric g-formula (20), a generalization of
standardization to time-varying exposures and confound-
ers, to estimate the 10-year risk of death under each physical
activity strategy. The g-formula appropriately handles treatment-
confounder feedback that occurs when the measured time-varying confounders are affected by prior physical activity and, under the assumptions of no unmeasured confounding and no model misspecification, validly estimates the risk had all eligible individuals adhered to the specified strategies (21).

The parametric g-formula has been previously applied to estimate the effects of lifestyle interventions on risk of coronary heart disease (22–24), type 2 diabetes (25), stroke (26), and asthma (27).

We adjusted for the following baseline covariates: age, parental history of myocardial infarction before age 60 years, primary treatment, clinical stage, Gleason grade, and prostate-specific antigen level at diagnosis, and prebaseline values of BMI, smoking history, and vigorous and moderate physical activity; and for the following time-varying covariates: vigorous and moderate physical activity, BMI, and the development of conditions excluded at baseline (functional impairment, metastasis, myocardial infarction, stroke, congestive heart failure, and amyotrophic lateral sclerosis). All models included indicator variables for period of follow-up, baseline covariates, and time-varying covariates measured in the previous questionnaire. See Web Appendices 1 and 2, Web Table 2, and Web

Figure 1. Selection of participants for a study of physical activity and survival among men with nonmetastatic prostate cancer, Health Professionals Follow-up Study, 2000–2014. Baseline refers to the return date of the first questionnaire after prostate cancer diagnosis. ALS, amyotrophic lateral sclerosis; CHF, congestive heart failure; MI, myocardial infarction.
Figures 1–4 for details on the g-formula, covariates, and models.

We compared the estimated 10-year risks had all eligible men followed each of the strategies with the risk under the observed levels of physical activity—that is, no intervention or “natural course”—via a risk ratio and a risk difference. The population attributable risk is 1 minus this risk ratio. We used nonparametric bootstrapping with 500 samples to obtain percentile-based 95% confidence intervals. We also generated adjusted survival curves and estimated the amount of survival time gained and time invested in physical activity over 10 years under each strategy.

**Sensitivity analyses for unmeasured confounding.** A key threat for the validity of our estimates is potential confounding by cancer severity or chronic disease if these conditions are severe enough to affect both physical activity and risk of death. The g-formula provides a natural way to partly address this problem by estimating the risk under physical activity interventions that are only applied at time t to persons who are sufficiently healthy at time t. Thus, in our main analysis, we considered potential interventions under which men were excused from adhering to the recommended physical activity levels after the development of functional impairment, metastasis, or a serious condition (myocardial infarction, stroke, congestive heart failure, or amyotrophic lateral sclerosis) that might limit participation. In sensitivity analyses, we expanded the definition of a serious condition to also include angina pectoris, pulmonary embolism, heart rhythm disturbance, diabetes mellitus, chronic renal failure, rheumatoid arthritis, gout, ulcerative colitis/Crohn disease, emphysema, Parkinson disease, and multiple sclerosis.

However, the validity of this approach relies critically on the assumption that the available data contain sufficient information to identify persons who are not healthy enough to maintain the recommended physical activity levels (23). We therefore also conducted an alternative analysis in which physical activity and covariate data were lagged by 2 years (23). In addition, we used a negative outcome control to detect potential unmeasured confounding by clinical disease (28). Specifically, we examined the effect of physical activity on questionnaire nonresponse. We selected questionnaire nonresponse as an alternative outcome not directly affected by physical activity but for which the effect of physical activity would be similarly confounded (i.e., by disease severity). If the estimated effect of physical activity on nonresponse were null, then it would be less likely that our findings for physical activity and death were confounded through the same pathways as this negative control.

**Other sensitivity analyses.** We assessed the robustness of our estimates to various analytical decisions. Specifically, we 1) altered the covariate order when modeling the joint distribution of time-varying covariates reported in the same questionnaire, 2) defined different functional forms for the covariates, 3) additionally adjusted for a baseline healthy behavior score (regular multivitamin use, routine physical examinations, and screening via rectal examination, sigmoidoscopy, or colonoscopy), 4) kept the total amount of time devoted to vigorous and moderate physical activity constant (something not implied by the guidelines) by replacing vigorous activity with moderate activity and vice versa, 5) removed the cardiovascular baseline exclusions and intervened on men who developed these conditions over time, and 6) repeated analyses assuming that the development of functional impairment, metastasis, and chronic diseases occurred 2 years before the questionnaire in which they were reported.

All analyses were conducted using SAS 9.3 (SAS Institute, Inc., Cary, North Carolina). The SAS GFORMULA macro and its documentation are available at http://www.hsph.harvard.edu/causal/software.

**RESULTS**

Table 1 shows baseline characteristics of the 2,299 eligible men. The mean age was 69 years, the mean BMI was 26, the mean level of vigorous activity was 1.8 hours/week, and the

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>68.9 (6.3)</td>
<td>93</td>
</tr>
<tr>
<td>Body mass index</td>
<td>26.0 (3.2)</td>
<td>2</td>
</tr>
<tr>
<td>Current physical activity, hours/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vigorous</td>
<td>1.8 (2.7)</td>
<td>41</td>
</tr>
<tr>
<td>Moderate</td>
<td>7.1 (7.5)</td>
<td>41</td>
</tr>
<tr>
<td>Caucasian race/ethnicity</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Parental history of MI (at age ≤60 years)</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>History of smoking before diagnosis</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>PSA level at diagnosis, ng/mL</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>&lt;4</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>4–10</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>≥10</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Clinical stage at diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>T3, T4, N1/M0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Gleason grade at diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;7</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>&gt;7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Primary treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radical prostatectomy</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Radiation therapy</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Hormones</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Watchful waiting</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: MI, myocardial infarction; PSA, prostate-specific antigen; SD, standard deviation.

a Baseline refers to the return date of the first questionnaire after prostate cancer diagnosis.

b Body mass index was calculated as weight (kg)/height (m)^2.

c Tumor-node-metastasis classification system (12).
mean level of moderate activity was 7.1 hours/week. Most men had clinically localized (T1/T2) and low-grade tumors. Of these men, 250 died within 10 years of prostate cancer diagnosis and 147 did not complete the full follow-up. Table 2 shows the estimated 10-year risk of death under strategies that excuse men from following the recommended physical activity levels after development of functional impairment, metastasis, or a serious cardiovascular or neurological condition. Compared with the 15.4% risk under no intervention, the risks were 13.0% (95% confidence interval: 10.9, 15.4) and 11.1% (95% CI: 8.7, 14.1) for ≥1.25 hours/week and ≥2.5 hours/week of vigorous activity and 13.9% (95% CI: 10.6, 14.7) for ≥2.5 hours/week of moderate activity, respectively. The proportion of men intervened on under these strategies is shown in Web Table 3. P values were calculated at the request of the Editor (Web Tables 4 and 5).

Figure 2 shows the estimated survival curves under these strategies. We estimated that average survival was 1.3 months, 2.4 months, 0.9 months, and 1.7 months longer for ≥1.25 and ≥2.5 hours/week of vigorous activity and ≥2.5 and ≥5 hours/week of moderate activity, respectively, compared with no intervention (Web Table 6). These averages should be interpreted in the context that most men did not die during follow-up. We estimated that an additional 0.7 months, 1.5 months, 0.5 months, and 1.2 months of physical activity were invested in the respective strategies over 10 years, compared with no intervention (Web Table 6).

Expanding the set of chronic conditions that would excuse men from following the recommended physical activity levels attenuated the differences across interventions, although risk was still lower with increasing duration and intensity of physical activity (Table 3).

The estimated risk ratio for questionnaire nonresponse, our negative outcome control, was close to null (Web Table 7). Estimates were similar when we altered the order in which we modeled time-varying covariates reported in the same questionnaire, defined different functional forms for covariates, additionally adjusted for a baseline healthy behavior score, and included persons with a recent cardiovascular event or congestive heart failure in our study population (data not shown). Estimates were also similar when we assumed that the development of functional impairment, metastasis, and chronic diseases occurred 2 years before the questionnaire in which these conditions were reported (Web Table 8). Risk ratio estimates were attenuated in analyses that kept the total time devoted to vigorous and moderate physical activity constant and that imposed a 2-year lag (Web Tables 9 and 10); for 1.25 hours/week of vigorous activity, risk ratios were 0.87 and 0.85, respectively, compared with 0.84 in Table 2; for 2.5 hours/week of moderate activity, risk ratios were 0.93 and 0.95, respectively, compared with 0.90 in Table 2.

**DISCUSSION**

In the United States, there are an estimated 160,000 men diagnosed with prostate cancer each year and 3 million survivors alive today (29). Most of these men will die from causes other than their cancer (30). Given the limited feasibility of a randomized trial of physical activity and long-term survival among men with nonmetastatic prostate cancer, we used observational data to generate the first (to our knowledge) survival estimates for adhering to current physical activity guidelines after diagnosis until it becomes impractical to do so. Our results suggest that 13%–16% of deaths in our population...
Observational studies have compared relative rates of death among prostate cancer survivors reporting different levels of postdiagnosis physical activity (1–4). Three studies of US and Swedish men included persons with nonmetastatic or localized prostate cancer (1–3), and 1 study of Canadian men also included those with metastatic disease (4). Physical activity exposure has been defined as total activity (1–4), walking or biking (2, 3), household work (2), “exercising” (2), vigorous activity (1, 4), or nonvigorously activity (1), in units of MET-hours/week or time/week. Two studies adjusted for prediagnosis activity (1, 4), and 2 did not (2, 3). Two presented unadjusted Kaplan-Meier curves (2, 4).

Of the studies with an exposure definition most comparable to ours, Friedenreich et al. (4), in the Prostate Cohort Study, reported an all-cause mortality hazard ratio of 0.65 (95% CI: 0.46, 0.92) for >3.5 hours/week of postdiagnosis vigorous activity versus none among Canadian men, which is consistent with our findings for a similar contrast. In a previous analysis in the HPFS, Richman et al. (1) reported an all-cause mortality hazard ratio of 0.51 (95% CI: 0.36, 0.72) for ≥3 hours/week of postdiagnosis vigorous activity versus <1 hour/week, after adjustment for mortality risk factors and prediagnosis physical activity. This is similar to our point estimate from conventional pooled logistic models (Web Table 11) but more protective than our g-formula estimate (risk ratio = 0.66) comparing realistic strategies in which men follow the guidelines until the development of conditions limiting physical ability (Web Table 12). Also, unlike previous analyses, our approach allows us to report adjusted estimates of absolute risk and population attributable risk for physical activity that is sustained until impractical.

Our novel approach has several additional strengths. We designed our study by explicitly specifying the protocol of the target trial and its observational emulation, an approach promoted in observational studies to reduce risk of bias and improve data interpretation (11). We analyzed our study by applying the parametric g-formula to high-quality observational data, allowing us to estimate clinically meaningful absolute risks under long-term strategies that realistically depend on evolving risk factors. Lastly, we performed several analyses to address potential reverse causation that leveraged stopping criteria as a natural way to address confounding by clinical disease as well as negative outcome controls as a tool for detecting bias in observational studies (28).

As with all analyses of observational data, the validity of our estimates relies on assumptions of no unmeasured confounding, measurement error, or model misspecification. Though we adjusted for many common causes of physical activity and death, the possibility of unmeasured or residual confounding by disease severity cannot be excluded. However, evidence for a protective effect of physical activity was robust to sensitivity analyses for reverse causation. We relied on self-reported information, which is subject to measurement error. However, previous validation studies in the HPFS have shown that self-reported weight and technician-measured weight are highly correlated (Pearson’s r = 0.97) and that the physical activity questionnaire is reproducible and provides a useful measure of average weekly activity (17, 19). Finally, our estimates may not be generalizable to other populations with different distributions of physical activity, risk factors, or effect modifiers.
In summary, we estimated that these US men with nonmetastatic prostate cancer would have experienced clinically meaningful reductions in mortality had they all followed current physical activity recommendations until they developed conditions limiting their physical ability. These findings may help guide clinical recommendations for prostate cancer patients and the design of future randomized trials.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Barbara A. Dickerman, Edward Giovannucci, Claire H. Pernar, Lorelei A. Mucci, Miguel A. Hernán); Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Edward Giovannucci); Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts (Edward Giovannucci, Lorelei A. Mucci); Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Miguel A. Hernán); and Harvard-MIT Division of Health Sciences and Technology, Boston, Massachusetts (Miguel A. Hernán).

B.A.D. was supported by an ASISA Fellowship (ASISA, Madrid, Spain). B.A.D. and C.H.P. were supported by National Cancer Institute grant T32 CA009001. L.A.M. was supported by a Young Investigator Award from the Prostate Cancer Foundation. M.A.H. was supported by National Cancer Institute grant P01 CA134294. The Health Professionals Follow-up Study is supported by National Cancer Institute grant U01 CA167552.

We thank the participants and staff of the Health Professionals Follow-up Study for their valuable contributions, as well as the following state cancer registries for their help: Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Nebraska, New Hampshire, New Jersey, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Virginia, Washington, and Wyoming.

Conflict of interest: none declared.

REFERENCES
